



ELSEVIER

Journal of Computational and Applied Mathematics 103 (1999) 99–114

JOURNAL OF
COMPUTATIONAL AND
APPLIED MATHEMATICS

A posteriori estimators for nonlinear elliptic partial differential equations

Felix Christian, Guimarães Santos *

Computational Mechanics Group, Department of Mechanical Engineering, UFPE, Recife PE 50740-530, Brazil

Received 9 October 1997; received in revised form 3 July 1998

Abstract

Many works have reported results concerning the mathematical analysis of the performance of a posteriori error estimators for the approximation error of finite element discrete solutions to linear elliptic partial differential equations. For each estimator there is a set of restrictions defined in such a way that the analysis of its performance is made possible. Usually, the available estimators may be classified into two types, i.e., the implicit estimators (based on the solution of a local problem) and the explicit estimators (based on some suitable norm of the residual in a dual space). Regarding the performance, an estimator is called asymptotically exact if it is a higher-order perturbation of a norm of the exact error. Nowadays, one may say that there is a larger understanding about the behavior of estimators for linear problems than for nonlinear problems. The situation is even worse when the nonlinearities involve the highest derivatives occurring in the PDE being considered (strongly nonlinear PDEs). In this work we establish conditions under which those estimators, originally developed for linear problems, may be used for strongly nonlinear problems, and how that could be done. We also show that, under some suitable hypothesis, the estimators will be asymptotically exact, whenever they are asymptotically exact for linear problems. Those results allow anyone to use the knowledge about estimators developed for linear problems in order to build new reliable and robust estimators for nonlinear problems. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Finite Element Method; Nonlinear PDEs; A posteriori error estimators

1. Introduction

This work deals with the relationship between the approximation error of finite element solutions to strongly nonlinear elliptic partial differential equations in some norms, with the error estimators computed for some suitably defined linear elliptic partial differential equations. It will be proved in what follows that, provided the problem data are smooth, it is possible to build linear elliptic

* Corresponding author. E-mail: fcgs@dmate.ufpe.br.

problems such that their finite element error is asymptotically equal to the finite element error for the nonlinear problem, and, provided the estimator being used is asymptotically exact for smooth linear problems, then, it will also be asymptotically exact for that auxiliary linear problem and, consequently for the original nonlinear problem. This work is more precisely developed and more general than the theory presented in [12]. The estimators considered are in a very large class, including virtually all implicit estimators, i.e., those estimators computed through the solution of local elliptic problems (either patchwise or elementwise). This is important in the sense that one may use estimators for nonlinear problems in the same fashion as it is done for linear problems, with the understanding that the same advantages and disadvantages of any particular estimator originally developed for linear problems will occur when used for nonlinear problems.

Works concerning error estimators for nonlinear problems are not equal in number and accomplishments to those concerning linear problems. In this work we do not have the intention of reviewing the literature in that field, but it is relevant to cite some important work, in which either similar or different strategies were used, when compared with our approach. The idea of computing estimators through linear problems can be traced back to the abstract works of Krasnosel'skii and collaborators [9], and, when related to a formal framework of the finite element method, to the works of Babuška and Rheinboldt [2, 11]. More recently, regarding strongly nonlinear elliptic partial differential equations with quadratic growth, Tsuchiya has also cited the relation between error estimation for linear problems and for nonlinear problems [13] (see also [10]). Verfürth has developed a method of estimating the norm of the residual, which is equivalent to the error of the nonlinear problem [14]. The disadvantage of the strategy related to estimating the residual (explicit error estimators), is that the estimator can only be proved to be equivalent to the error, therefore, including some multiplying constants which may be either small or large depending on the problem.

In this work we deal only with regular points, because it allows for a more direct approach, making it easier to convey the main ideas. Extensive numerical experiments will be presented in [12], including examples with known solutions.

Let $F : W_0^{1,p_1} \times \mathbb{R}^m \rightarrow W^{-1,p'_2}$ be given and consider the following problem:

Pr 1. Find $(u_0, \lambda_0) \in (W_0^{1,p_1} \times \mathbb{R}^m)$ such that

$$F(u_0, \lambda_0) = 0 \quad \text{on } \Omega,$$

where $\Omega \subset \mathbb{R}^2$ is open and bounded. Here

$$F(u, \lambda) = -\nabla \cdot [\mathbf{a}(\nabla u, u, \lambda, x)] + \mathbf{b}(\nabla u, u, \lambda, x) + \mathbf{c}(u, \lambda, x) - \mathbf{f}(\lambda, x);$$

where $u : \Omega \rightarrow \mathbb{R}$; $\mathbf{a} : \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$; $\mathbf{b} : \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^2 \rightarrow \mathbb{R}$; $\mathbf{c} : \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^2 \rightarrow \mathbb{R}$; $\mathbf{f} : \mathbb{R}^m \times \mathbb{R}^2 \rightarrow \mathbb{R}$ are given smooth enough functions.

In this work we are interested in the a posteriori numerical analysis, so we are going to assume that the following hypothesis holds

Hypothesis 1.1. *There exists a nonempty set $A \subset \mathbb{R}^m$ such that, for all $\lambda_0 \in A$, Pr.1 has at least one solution point $u_0(\lambda_0)$ in some given admissible closed convex set $\mathcal{K} \subset W_0^{1,p_1}$.*

The discrete problem is set as

DPr. 1. Find $(u_h, \lambda_h) \in (S^h(\tau_h, p, \Omega) \cap \mathcal{X}) \times \mathbb{R}^m$ such that

$$\langle F(u_h, \lambda_h), v_h \rangle = 0, \quad \text{for all } v_h \in S^h(\tau_h, p, \Omega) \cap W_d^{1,p_2}.$$

The main issue now is to establish some restrictions on the differential equations we are going to deal with. Actually, there are further issues which will not be covered here, but the reader will find them in [3], where a more complete description of the hypothesis will be found. For a more detailed analysis on the differentiability structure required in the hypothesis stated below see [3]; for the existence and convergence results we refer to [3].

Let us be specific and assume that $F : W_0^{1,p_1} \times \mathbb{R}^m \rightarrow W^{-1,p'_2}$, $1 < p_1 < \infty$, $1 < p_2$, is defined by

$$F(u, \lambda) = Q(u, \lambda) - R(u, \lambda) - f(\lambda). \quad (1)$$

Here $Q(\cdot, \lambda) : W_0^{1,p_1} \rightarrow W^{-1,p'_2}$ is an isomorphism and a strongly nonlinear operator for all $\lambda \in \mathbb{R}^m$; $R : W_0^{1,p_1} \times \mathbb{R}^m \rightarrow W^{-1,p'_2}$ is a compact and smooth nonlinear operator, and $f(\lambda) \in W^{-1,p'_2}$.

Also, assume

Hypothesis 1.2. $F : W_0^{1,p_1} \times \mathbb{R}^m \rightarrow W^{-1,p'_2}$ satisfies the following properties:

- (i) F is a Θ -Fredholm operator of index $i(F) = m$ from $\Theta = (W_0^{1,p_1}, H_0^1, W_0^{1,\infty})$ into $\Theta^* = (W^{-1,p'_2}, H^{-1}, W^{-1,\infty})$.
- (ii) The extension

$$D_u Q_H(w, \lambda) : H_0^1 \rightarrow H^{-1}$$

is a coercive and bounded linear operator for all $w \in W^{1,\infty}$ and $\lambda \in \mathbb{R}^m$, with the constants of boundedness and coercivity being bounded uniformly away from ∞ and 0, respectively, in bounded sets of $(w, \lambda) \in W^{1,\infty} \times \mathbb{R}^m$. Furthermore, its coefficients are in L^∞ .

- (iii) For all $u_1, u_2 \in W_0^{1,\infty}$ and $\lambda \in \mathbb{R}^m$, there exists $C = C(\|u_1\|_{W^{1,\infty}}, \|u_2\|_{W^{1,\infty}}, |\lambda|)$, such that

$$\|D_u Q_H(u_1, \lambda) - D_u Q_H(u_2, \lambda)\|_{\mathcal{L}(H_0^1, H^{-1})} \leq C \|u_1 - u_2\|_{W^{1,\infty}}$$

and

$$\|D_u R(u_1, \lambda) - D_u R(u_2, \lambda)\|_{\mathcal{L}(H_0^1, H^{-1})} \leq C \|u_1 - u_2\|_{W^{1,\infty}}.$$

Furthermore, all the coefficients of $D_u F(u_0, \lambda_0)$ are as smooth as the gradient of u_0 .

- (iv) For all $(u, \lambda) \in W_0^{1,\infty} \times \mathbb{R}^m$, all the existing derivatives of F with respect to the function and the parameter at (u, λ) are Hölder-continuous with respect to λ . Moreover, the existing derivatives of F with respect to the parameter are in $W^{-1,p}$, for all needed values of p , and are Hölder-continuous with respect to the function.
- (v) For all $(u, \lambda) \in W_0^{1,\infty} \times \mathbb{R}^m$, the linear operator $D_u R(u, \lambda) : W_0^{1,p} \rightarrow W^{-1,p'}$ is a compact operator for all p , $r' \leq p \leq r$, with $r > 2$ as large as needed.

The main results in this paper depend on some properties of linearized operators. So we define:

Hypothesis 1.3. (Continuous inf-sup condition). Let $\infty > r \geq 2$ be any number, $1/r + (1/r') = 1$, and $B : H_0^1 \times H_0^1 \rightarrow \mathbb{R}$ be as above. Then, $B : W_0^{1,t} \times W_0^{1,t'} \rightarrow \mathbb{R}$ is bounded and satisfies

$$\inf_{u \in W_0^{1,t}} \left\{ \sup_{v \in W_0^{1,t'}} \left\{ \frac{B(u,v)}{\|u\|_{W^{1,t}} \|v\|_{W^{1,t'}}} \right\} \right\} \geq \gamma > 0,$$

$$\sup_{u \in W_0^{1,t}} \{B(u,v)\} > 0 \quad \text{for all } v \in W_0^{1,t'},$$

for all $t \in [r', r]$.

Hypothesis 1.4. (Discrete inf-sup condition). Let $\infty > r \geq 2$ be any number, $1/r + (1/r') = 1$, and $B : H_0^1 \times H_0^1 \rightarrow \mathbb{R}$ be as above. Then, $B : W_0^{1,t} \times W^{1,t'} \rightarrow \mathbb{R}$ is bounded and satisfies

$$\inf_{u_h \in S^h} \left\{ \sup_{v_h \in S^h} \left\{ \frac{B(u_h, v_h)}{\|u_h\|_{W^{1,t}} \|v_h\|_{W^{1,t'}}} \right\} \right\} \geq \theta > 0,$$

$$\sup_{u_h \in S^h} \{B(u_h, v_h)\} > 0 \quad \text{for all } v_h \in S^h,$$

for all $t \in [r', r]$. Also, $\theta \neq \theta(h)$.

First, for any $q \in [r', r]$, $h > 0$, $w_h \in S^h$ and $w \in W^{1,\infty}$, define

$$\beta_h = \beta_h(h, q, w) = \max \left\{ \inf_{v_h \in S^h} \left\{ h^{-n/q} \|w - v_h\|_{W^{1,q}} \right\}, \inf_{v_h \in S^h} \{\|w - v_h\|_{W^{1,\infty}}\} \right\}.$$

The following results are statements concerning existence and convergence of discrete solutions to DPr.1. For that, let $K = D_u(u_0, \lambda_0)$ and $B(\cdot, \cdot) = \langle K(\cdot), (\cdot) \rangle$.

Theorem 1.5. Let $F : W_0^{1,p_1} \times \mathcal{R}^m \rightarrow W^{-1,p_2'}$ satisfy Hypothesis 1.2. Let the linear operator $K : H_0^1 \rightarrow H^{-1}$ be defined as above and let the bilinear form $B : H_0^1 \times H_0^1 \rightarrow \mathbb{R}$ be defined based on K . Let $\Omega \in \mathcal{D}^t$, for some $t > 2$. Let (u_0, λ_0) be a strong regular solution point to Pr.1, such that $u_0 \in W_0^{1,p_1} \cap W^{1+\xi,p}$, with $\xi > n/p$.

Then, there is an $r > 2$, such that, if one can find $q \in [r', r]$, satisfying

$$\beta_h(h, q, u_0) \rightarrow 0,$$

one can also find $h_0 > 0$, and $\delta > 0$, such that, for all $0 < h \leq h_0$, there exists a unique $u_h \in \tilde{W}_0^{1,q} \cap B_\delta(\bar{u}_h)$, such that (u_h, λ_0) solves DPr.1. Here the sequence $\{\bar{u}_h\}_{h \rightarrow 0} \in S^h$ is to be suitably chosen.

Proof. This result is just a particular case (for regular solution points) of a more general theorem presented in [3]. \square

Remark. The sequence $\{\bar{u}_h\}_{h \rightarrow 0}$ in the above theorem is to be chosen such that the objectives of the analysis are met [3, 12].

Corollary 1.6. Let the hypotheses of Theorem 1.5 be true. Then, for all $h > 0$ small enough,

$$\|u_h - u_0\|_{W^{1,s}} \leq C \|\bar{u}_h - u_0\|_{W^{1,s}},$$

for all $s \in [r', r]$, such that $\beta_h(h, s, u_0) \rightarrow 0$, as $h \rightarrow 0$;

$$\|u_h - u_0\|_{W^{1,s}} \leq C[h^{n/s - (n/r)} \|\bar{u}_h - u_0\|_{W^{1,r}} + \|\bar{u}_h - u_0\|_{W^{1,s}}],$$

for all $s \in [r, \infty]$, where $\{\bar{u}_h\}_{h \rightarrow 0}$ is a sequence chosen as in Theorem 1.5. Here $C = C(u_0, \lambda_0)$ and $r > 2$ is as obtained in Theorem 1.5.

Furthermore, taking $\bar{u}_h = P_h u_0$, and for all $s \in [r', \infty]$, such that $\xi > n/\min\{s, p\}$, the above inequalities imply that

$$\|u_h - u_0\|_{W^{1,s}} \leq Ch^\mu \|u_0\|_{W^{1+\xi, p}},$$

where, $\mu_1 = \min\{0, n/s - (n/r)\}$ and $\mu = \min\{q, \mu_2\}$, where $q \geq 1$ is the polynomial order of approximation of the shape functions in each finite element, and $\mu_2 = \xi + (n/s) - n/\min\{r, p\}$ if $s \geq r$ and $\mu_2 = \xi + (n/s) - n/\min\{s, p\}$ if $s < r$; and, $C \neq C(u_0, h)$.

Proof. This result is a particular case (for regular solution points) of a more general theorem presented in [3].

2. A posteriori estimators

In this section we develop a procedure for relating computable a posteriori error estimators for a suitably defined auxiliary linear problem with the exact error (in the norm of $W^{1,s}$, $s \in [r', r]$, $r \geq 2$) for the nonlinear problem (between a given solution to Pr.1 and the corresponding discrete solution to DPr.1). A large class of estimators will be considered, namely, implicit estimators, obtained through a solution of a suitably defined local problem and defined either elementwise or patchwise. In order to make the procedures clear, we will consider only strong regular solution points. The procedures regarding simple turning points will be presented in later works.

A first linear auxiliary problem will be defined by a bilinear form $B_1 : H_0^1 \times H_0^1 \rightarrow \mathbb{R}$ and a right-hand side f_1 . Similarly, a second linear auxiliary problem will be defined by $B_0 : H_0^1 \times H_0^1 \rightarrow \mathbb{R}$ and f_0 . For the exact, discrete and error equations of both problems we refer to LP.0, DLP.0 and Er.0 below, respectively.

In what follows, the expression $\omega \cap T \neq \emptyset$ will mean that the interior of the region defined by ω has an empty intersection with the region defined by T .

Definition 2.1. Let a mesh τ_h be given. Suppose that a way of building a set of patches ω by making union of adjacent elements $T \in \tau_h$, such that the union of all patches covers Ω , is given. Let \mathcal{V}_h be that set. Define

(a) The trial space

$$W_\omega = \text{span}\{\psi_j^\omega\}_{j=1}^{j=k(\omega)},$$

and the test space

$$V_\omega = \text{span}\{\phi_j^\omega\}_{j=1}^{j=k(\omega)},$$

defined over each corresponding patch $\omega \in \mathcal{V}_h$.

(b) The spaces (defined elementwise), for each $T \in \tau_h$,

$$Y_T = \text{span}\{\psi_j^\omega|_T, \quad j = 1, \dots, k(\omega), \text{ for all } \omega, \text{ such that } \omega \cap T \neq \emptyset\},$$

$$Z_T = \text{span}\{\phi_j^\omega|_T, \quad j = 1, \dots, k(\omega), \text{ for all } \omega, \text{ such that } \omega \cap T \neq \emptyset\}.$$

It is clearly seen that there exist a decomposition $w_h^T = \sum_{\omega \cap T \neq \emptyset} [w_h^\omega]$, for all $w_h^T \in Y_T$, where $w_h^\omega \in W_\omega$ with $\omega \cap T \neq \emptyset$. Similarly, there exists a decomposition $v_h^T = \sum_{\omega \cap T \neq \emptyset} [v_h^\omega]$, for all $v_h^T \in Z_T$, where $w_h^\omega \in W_\omega$.

Hypothesis 2.2. Let a bilinear form $B(\cdot, \cdot) : W_0^{1,s} \times W^{1,s'} \rightarrow \mathbb{R}$ be given, for all $s \in [r', r]$, where $r \in [2, \infty)$ is given. Let a mesh τ_h and the set of patches \mathcal{V}_h be given as defined above. Then,

(a) There exists a real number $\gamma > 0$, such that, for each $T \in \tau_h$, and for all $w_h \in Y_T$,

$$\gamma \|w_h\|_{W^{1,s}(T)} \leq \sup_{v_h^T \in Z_T} \left\{ \sum_{\omega \cap T \neq \emptyset} \left[\frac{B_\omega(w_h^\omega, v_h^\omega)}{\|v_h^\omega\|_{W^{1,s'}(\omega)}} \right] \right\},$$

where $\gamma \neq \gamma(h)$ does not depend either on $T \in \tau_h$, nor on \mathcal{V}_h . $B_\omega(\cdot, \cdot)$ means the restriction of B to the patch ω . Note that we are using the decomposition of $v_h^T \in Z_T$ described just above.

(b) Let $R \in W^{-1,r}$ be given, and consider $R|_\omega$ as being a suitably defined restriction of R to ω , for all $\omega \in \mathcal{V}_h$. Then, there exist positive constants C_1 and C_2 , not depending either on the τ_h nor on \mathcal{V}_h , such that

$$C_1 \|R\|_{W^{-1,s}(\Omega)} \leq \left[\sum_{T \in \tau_h} \left(\sum_{\omega \cap T \neq \emptyset} \|R|_\omega\|_{W^{-1,s}(\omega)} \right) \right]^{1/s} \leq C_2 \|R\|_{W^{-1,s}(\Omega)}.$$

Remark. Hypothesis 2.2 means that the given bilinear form B is patchwise elliptic and the patches do not overlap too much, destroying the stability of the sum of quantities defined patchwise.

We now define the class of implicit estimators.

Definition 2.3. (Implicit estimators). Let a mesh τ_h and a set of patches \mathcal{V}_h be given. Let a bilinear form $B(\cdot, \cdot)$ be given, which satisfies Hypotheses 1.3, 1.4, and 2.2 for some $r \in [2, \infty)$ and spaces $\{V_\omega\}_{\omega \in \mathcal{V}_h}$ and $\{W_\omega\}_{\omega \in \mathcal{V}_h}$. Let $f \in W^{-1,r}$ be given, and define $w_0 \in W_0^{1,r}$, $w_h \in S^h(\tau_h)$ and $e \in W_0^{1,r}$ to be the solutions of LP.0, DLP.0 and Er.0, respectively. For each $T \in \tau_h$, define $\xi_T(x) \in Y_T$ as

$\xi_T(x) = \sum_{\omega \cap T \neq \emptyset} [\sum_{j=1, \dots, k(\omega)} C_j^\omega \psi_j^\omega(x)]$, for $x \in T$, where $\{\{C_j^\omega\}_{j=1}^{j=k(\omega)}\}_{\omega \cap T \neq \emptyset}$ are constants, which are obtained by finding $\xi^\omega \in W_\omega$, such that

$$B_\omega(\xi^\omega, \phi_j^\omega) = \langle R_{h\omega}, \phi_j^\omega \rangle \quad j = 1, \dots, k(\omega) \quad \text{for all } \omega \in \mathcal{V}_h,$$

where $R_{h\omega} = R_h|_\omega$ is the restriction of the residual $R_h \in W^{-1,r}$ to the patch $\omega \in \mathcal{V}_h$. Thus, $\xi_T(x)$ is computed by

$$\xi_T(x) = \sum_{\omega \cap T \neq \emptyset} [\xi^\omega(x)],$$

for all $x \in T$. The restricted residual R_ω , the trial and the test spaces should be such that the above problem has a unique solution.

For some given $s \geq 1$, set

$$\eta_T(s) = \|\xi_T\|_{W^{1,s}(T)},$$

and

$$\eta(s) = \left\{ \sum_{T \in \tau_h} \eta_T^s \right\}^{1/s}.$$

The value η_T is called the elemental estimator for $T \in \tau_h$ (indicator) and η is the (global) estimator.

The auxiliary linear problems are to be defined as follows.

LP.0. Find $w \in W_0^{1,s}$, such that

$$B(w, v) = \langle f, v \rangle \quad \text{for all } v \in W_0^{1,s'}.$$

DLP.0. Find $w_h \in S^h$, such that

$$B(w_h, v_h) = \langle f, v_h \rangle \quad \text{for all } v_h \in S^h,$$

where $f \in W^{-1,s'}$. Defining the error by $e = w - w_h$, the error equations for the above problems are given by

Er.0. Find $e \in W_0^{1,s}$, such that

$$B(e, v) = \langle R_h, v \rangle = B(w_h, v) - \langle f, v \rangle \quad \text{for all } v \in W_0^{1,s'}.$$

The following result shows that the implicit estimators change at most linearly with perturbations in the coefficients of the operators and on the right-hand side.

Theorem 2.4. Let the bilinear forms $B_0(\dots) : H_0^1 \times H_0^1 \longrightarrow \mathbb{R}$ and $B_1(\dots) : H_0^1 \times H_0^1 \longrightarrow \mathbb{R}$ be given. Let both bilinear forms satisfy Hypotheses 1.3, 1.4 and 2.2 for some $r \in [2, \infty)$ and spaces $\{W_\omega\}_{\omega \in \mathcal{V}_h}$ and $\{V_\omega\}_{\omega \in \mathcal{V}_h}$. Let $f_0, f_1 \in W^{-1,r}$ be given as the right-hand sides for B_0 and B_1 , respectively. Let w_0 and $w_1 \in W_0^{1,r}$ be solutions of LP.0, $w_{0h}, w_{1h} \in S^h$ solutions of DLP.0 and $R_{0h}, R_{1h} \in W^{-1,r}$ the residuals, all related to B_0 and B_1 , respectively. If η_0 and η_1 are implicit estimators related to B_0, f_0 and B_1, f_1 , respectively, then, there exists a constant C , which depends only on the $L^\infty(\Omega)$ -norm of the coefficients of both bilinear norms, such that, for each $s \in [r', r]$,

$$|\eta_0(s) - \eta_1(s)| \leq C[\|\Delta B\|_{L^\infty} \eta_0 + \|\Delta R_h\|_{W^{-1,s}(\Omega)}].$$

For the above, $\|\Delta B\|_{L^\infty}$ means the $L^\infty(\Omega)$ -norm of the difference between the respective coefficients of B_0 and B_1 ; and ΔR_h is the difference between the residuals R_{0h} and R_{1h} .

Proof. Let $T \in \tau_h$ be given. Let $\xi_{0T}, \xi_{1T} \in Y_T$ be as in Definition 2.3, related to B_0 and B_1 , respectively. Set $\Delta \xi_T = \xi_{0T} - \xi_{1T}$, and $\Delta B(\cdot, \cdot) = B_1(\cdot, \cdot) - B_0(\cdot, \cdot)$. Then

$$\sum_{\omega \cap T \neq \emptyset} [\langle \Delta R_{h\omega}, \phi^\omega \rangle_\omega] = \sum_{\omega \cap T \neq \emptyset} [B_{0\omega}(\Delta \xi^\omega, \phi^\omega) + \Delta B_\omega(\xi_0^\omega, \phi^\omega)], \quad (2)$$

for all $\phi^\omega \in V_\omega$, $\omega \cap T \neq \emptyset$. Then, by Hypothesis 2.2 and for each $s \in [r', r]$, we obtain

$$\begin{aligned} \gamma \|\Delta \xi_T\|_{W^{1,s}(T)} &\leq \sup_{v_h^\omega \in Z_T} \left\{ \sum_{\omega \cap T \neq \emptyset} \left[\frac{B_{1\omega}(\Delta \xi_h^\omega, v_h^\omega)}{\|v_h^\omega\|_{W^{1,s'}(\omega)}} \right] \right\} \\ &\leq \sup_{v_h^\omega \in Z_T} \left\{ \sum_{\omega \cap T \neq \emptyset} \left[\left| \frac{\langle \Delta R_{h\omega}, v_h^\omega \rangle_\omega}{\|v_h^\omega\|_{W^{1,s'}(\omega)}} \right| + \left| \frac{\Delta B_\omega(\xi_0^\omega, v_h^\omega)}{\|v_h^\omega\|_{W^{1,s'}(\omega)}} \right| \right] \right\}. \end{aligned}$$

So, there is $C = C(s, \gamma)$, such that

$$\|\Delta \xi_T\|_{W^{1,s}(T)}^s \leq C \sum_{\omega \cap T \neq \emptyset} [\|\Delta R_{h\omega}\|_{W^{-1,s}(\omega)}^s + \|\Delta B_\omega\|_{L^\infty(\omega)}^s \|\xi_0^\omega\|_{W^{1,s}(\omega)}^s].$$

By adding up over all elements of τ_h and using the summation properties of the residuals stated in Definition 2.3, we get

$$\sum_{T \in \tau_h} \|\Delta \xi_T\|_{W^{1,s}(T)}^r \leq C \left[\|\Delta R_h\|_{W^{-1,s}(\Omega)}^s + \|\Delta B\|_{L^\infty(\Omega)}^s \sum_{T \in \tau_h} \|\xi_{0T}\|_{W^{1,s}(T)}^s \right].$$

Taking the s th-root on both sides of the above expression and using Minkowskii's inequality yields

$$|\eta_0 - \eta_1| \leq C [\|\Delta R_h\|_{W^{-1,s}(\Omega)} + \|\Delta B\|_{L^\infty(\Omega)} \eta_0],$$

which immediately gives the desired inequality. \square

Now, let us be specific and introduce our two auxiliary linear problems, the first for theoretical purposes only and the second for the actual computation of the error estimator. As before, and for the rest of this paper, (u_0, λ_0) and (u_h, λ_0) will be the solution to Pr.1 and the solution to DPr.1, respectively. Next, let $F : W_0^{1,p_1} \rightarrow W^{-1,p'_2}$ be given and satisfy the hypothesis of Theorem 1.5. For any fixed $r \in [2, \infty)$, as close to 2 as needed, we take $s \in [r', r]$ and set $K_0, K_{0h}, K_h : W^{1,s} \rightarrow W^{-1,s}$; and $B_0, B_1 : W_0^{1,s} \times W_0^{1,s'} \rightarrow \mathbb{R}$ as

$$K_0 = D_u F(u_0, \lambda_0), \quad (3)$$

$$K_{0h} = \int_0^1 [D_u F(u_h + t(u_0 - u_h), \lambda_0)] dt, \quad (4)$$

$$K_h = D_u F(u_h, \lambda_0), \quad (5)$$

$$B_0(u, v) = \langle K_0 u, v \rangle \quad \text{for all } u \in W_0^{1,s} \text{ and } v \in W_0^{1,s'}, \quad (6)$$

$$B_1(u, v) = \langle K_h u, v \rangle \quad \text{for all } u \in W_0^{1,s} \text{ and } v \in W_0^{1,s'}. \quad (7)$$

Furthermore, set

$$e_h = u_0 - u_h, \quad (8)$$

$$\Delta K_0 = K_{0h} - K_0, \quad (9)$$

$$\Delta K_1 = K_0 - K_h, \quad (10)$$

$$b_0 = K_0 u_0. \quad (11)$$

The next lemma states some properties for the above operators

Lemma 2.5. *Let the hypothesis of Theorem 1.5 and Corollary 1.6 be satisfied. Then, for each $s \in [r', r]$:*

- (i) *There exists a small enough $h_0 > 0$, such that for all $0 < h < h_0$, there exists a constant $C \neq C(h)$, such that*

$$\|\Delta K_0\|_{\mathcal{L}(W_0^{1,s}, W^{-1,s})} \leq C \|e_h\|_{W^{1,\infty}} \leq Ch^{\xi - n/\min\{r,p\}}.$$

- (ii) *There exists a small enough $h_0 > 0$, such that for all $0 < h < h_0$, there exists a constant $C \neq C(h)$, such that*

$$\|\Delta K_1\|_{\mathcal{L}(W_0^{1,s}, W^{-1,s})} \leq C \|e_h\|_{W^{1,\infty}} \leq Ch^{\xi - n/\min\{r,p\}}.$$

- (iii) *There exists a small enough $h_0 > 0$, such that for all $0 < h < h_0$, B_0 and B_1 satisfy Hypotheses 1.3 and 1.4.*

Proof. Items (i) and (ii) follow directly from item (iii) of Hypothesis 1.2 and Corollary 1.6.

Item(iii) is a consequence of the following facts: (a) K_0 satisfies the inf-sup condition with $r=2$, since (u_0, λ_0) is a strong regular point; (b) K_0 is a compact perturbation of $D_u Q(u_0)$, which satisfies Hypotheses 1.3 and 1.4, for some $r > 2$ [3]; (c) from item (ii) of the current lemma, it follows that K_h converges uniformly to K_0 for all $s \in [r', r]$. Then the result follows [3]. \square

The two auxiliary problems will be defined by the bilinear form B_1 (computable) and B_0 (abstract), together with the right-hand sides

$$f_1 = -F(u_h, \lambda_0), \quad (12)$$

$$f_0 = b_0, \quad (13)$$

respectively. Let $w_0^h \in W_0^{1,s}$ and $w_{0h}^h \in S^h$ solve LP.0 and DLP.0 with $B \equiv B_1$ and $f \equiv f_1$, that is,

$$B_1(w_0^h, v) = \langle -F(u_h, \lambda_0), v \rangle \quad \text{for all } v \in W_0^{1,s'}, \quad (14)$$

$$B_1(w_{0h}^h, v_h) = \langle -F(u_h, \lambda_0), v_h \rangle \quad \text{for all } v_h \in S^h. \quad (15)$$

Since by definition, u_0 solves LP.0 with $B \equiv B_0$ and $f \equiv f_0$, let $u_{0h} \in S^h$ solve the corresponding discrete problem (DLP.0), i.e.,

$$B_0(u_{0h}, v_h) = \langle b_0, v_h \rangle \quad \text{for all } v_h \in S^h. \quad (16)$$

Next, let us define the error expressions

$$e_{0h} = u_0 - u_{0h},$$

$$e_{wh} = w_{0h} - w_{0h}^h,$$

which are solutions to Er.0 for the abstract and computable auxiliary problems, respectively.

The next lemma establishes some further results regarding the relationship between both linear problems.

Lemma 2.6. *Let the hypothesis of Theorem 1.5 be satisfied. Then, the following statements are true:*

- (i) $K_0(u_0 - u_h) = -F(u_h, \lambda_0) + \Delta K_0 e_h$.
- (ii) $K_0 u_0 = b_0 = -F(u_h, \lambda_0) + \Delta K_0 e_h + K_0 u_h$.
- (iii) $B_0(u_{0h} - u_h, v_h) = \Delta B(e_h, v_h) = \langle \Delta K_0 e_h, v_h \rangle$, for all $v_h \in S^h$.
- (iv) There exists a constant $C \neq C(h)$, such that

$$\|u_{0h} - u_h\|_{W^{1,s}} \leq C \|\Delta K_0\|_{\mathcal{L}(W_0^{1,s}, W^{-1,s})} \|e_h\|_{W^{1,s}}$$

for all $s \in [r', r]$

- (v) $w_{0h}^h = 0$.
- (vi) Set

$$R_{0h} = b_0 - K_0 u_{0h},$$

and

$$R_{1h} = -F(u_h, \lambda_0) - K_h w_{0h}^h = -F(u_h, \lambda_0).$$

Then,

$$\Delta R_h = R_{0h} - R_{1h} = \Delta K_0 e_h + K_0(u_h - u_{0h})$$

and, hence, there exists a constant $C \neq C(h)$, such that

$$\|\Delta R_h\|_{W^{-1,s}} \leq C \|\Delta K_0\|_{\mathcal{L}(W_0^{1,s}, W^{-1,s})} \|e_h\|_{W^{1,s}}.$$

Proof. (i) This result comes from the observation that $K_{0h} e_h = -F(u_h, \lambda_0)$. Then,

$$K_0 e_h = K_{0h} e_h + (K_0 - K_{0h}) e_h = -F(u_h, \lambda_0) + \Delta K_0 e_h.$$

(ii) This relation comes directly from the result in (i).

(iii) Since $\langle F(u_h, \lambda_0), v_h \rangle = 0$, and $B_0(u_0 - u_{0h}, v_h) = 0$ for all $v_h \in S^h$, and from (i) we obtain

$$\begin{aligned} B_0(u_{0h} - u_h, v_h) &= B_0(u_{0h} + (u_0 - u_{0h}) - u_h, v_h) = B_0(u_0 - u_h, v_h) \\ &= \langle -F(u_h) + \Delta K_0 e_h, v_h \rangle = \langle \Delta K_0 e_h, v_h \rangle. \end{aligned}$$

(iv) From Lemma 2.5, B_0 satisfies Hypothesis 1.4, for some $r > 2$, and, then, with the help of (iii) we get

$$\theta \|u_h - u_{0h}\|_{W^{1,s}} \leq \sup_{v_h \in S^h} \left\{ \frac{B_0(u_h - u_{0h}, v_h)}{\|v_h\|_{W^{1,s'}}} \right\} = \sup_{v_h \in S^h} \left\{ \frac{\langle \Delta K_0 e_h, v_h \rangle}{\|v_h\|_{W^{1,s'}}} \right\},$$

where $\theta \neq \theta(h)$ and $s \in [r', r]$. Thus, the result follows immediately.

(v) Again, by Lemma 2.5, B_1 satisfies Hypothesis 1.4 for the same $r > 2$ as in item (iii) above, and, then, since $\langle F(u_h, \lambda_0), v_h \rangle = 0$, for all $v_h \in S^h$, we find that $w_{0h}^h = 0$.

(vi) This statement follows directly from (ii) and (iv).

Remark. As we said in the introduction, the strategy of using auxiliary linear problems for estimating the finite element error for nonlinear problems is not new. Therefore, property (v) in the lemma above were already known, no matter what linear coercive operator is used. The other remaining properties are new, and they show important issues. For instance, item (iv) shows that the finite element solution to the nonlinear problem is a higher-order perturbation of a finite element solution to a smooth linear problem. This opens a clear space for further investigations on the relation between superconvergence properties of nonlinear problems and those for linear problems.

Now, define $\eta_0(s)$ and $\eta_1(s)$ as being the same a posteriori estimator (with respect to the norm of $W_0^{1,s}$, $s \in [r', r]$) applied for estimating the errors $e_{0h} = u_0 - u_{0h}$ and $e_{wh} = w_{0h} - w_{0h}^h = w_{0h}$, respectively. The next theorem will state that both estimators will give the same result, asymptotically speaking (that is, when $h \rightarrow 0$).

Theorem 2.7. Consider $F = Q - R - f : W_0^{1,p_1} \times \mathbb{R}^m \longrightarrow W^{-1,p'_2}$ satisfying all the hypotheses of Theorem 1.5. Let the bilinear forms $B_0(\cdot, \cdot) : W_0^{1,s} \times W_0^{1,s'} \longrightarrow \mathbb{R}$ and $B_1(\cdot, \cdot) : W_0^{1,s} \times W_0^{1,s'} \longrightarrow \mathbb{R}$ be defined as above, for all $s \in [r', r]$, satisfying the hypotheses of Theorem 2.4. Consider u_0 and w_{0h} as the solutions to LP.0, related to the bilinear forms B_0 and B_1 , and the right-hand sides f_0 and f_1 , respectively. Let u_{0h} and w_{0h}^h be the respective approximate solutions to DLP.0, and R_{0h} and R_{1h} be the corresponding residuals, as described in problem Er.0. Let $\eta_0(s)$ and $\eta_1(s)$ ($s \in [r', r]$) be the result of the same implicit a-posteriori estimator applied to the abstract and the computable problems, respectively. If the given estimator satisfies Hypothesis 2.2, Then,

$$|\eta_0(s) - \eta_1(s)| \leq C \|e_h\|_{W^{1,\infty}} (\|e_h\|_{W^{1,s}} + \eta_0).$$

Here, as defined in (8), $e_h = u_0 - u_h$ is the error for the nonlinear problem Pr.1, and $C \neq C(h)$.

Proof. From Theorem 2.4 and Lemma 2.6 we get

$$|\eta_0(s) - \eta_1(s)| \leq C (\|\Delta B\|_{L^\infty} \eta_0 + \|\Delta K_0\|_{\mathcal{L}(W_0^{1,s}, W^{-1,s})} \|e_h\|_{W^{1,s}}).$$

From Lemma 2.5 we obtain that

$$\|\Delta B\|_{L^\infty} \leq C \|e_h\|_{W^{1,\infty}},$$

and

$$\|\Delta K_0\|_{\mathcal{L}(W_0^{1,s}, W^{-1,s})} \leq C \|e_h\|_{W^{1,\infty}}.$$

The three relations above yield the desired result. \square

The lemma below shows that all errors (nonlinear problem and auxiliary problems) are higher-order perturbations of each other, and that the estimator is asymptotically the same when applied to both auxiliary problems.

Lemma 2.8. *Let the hypothesis of Theorem 2.7 be satisfied. Furthermore, let $u_0 \in W^{1+\xi,p}$, such that $\xi > n/\min\{r, p\}$. Then, for all $s \in [r', r]$,*

- (i) $\|e_h - e_{wh}\|_{W^{1,s}} \leq Ch^\varepsilon \|e_h\|_{W^{1,s}};$
 - (ii) $\|e_{wh} - e_{0h}\|_{W^{1,s}} \leq Ch^\varepsilon \|e_{wh}\|_{W^{1,s}};$
 - (iii) $|\eta_0(s) - \eta_1(s)| \leq Ch^\varepsilon (\|e_{0h}\|_{W^{1,s}} + \eta_0(s)),$
- where $\varepsilon > 0$.

Proof. (i) This inequality follows from

$$\begin{aligned} \gamma \|e_h - e_{wh}\|_{W^{1,s}} &\leq \sup_{v \in W^{1,s'}} \left\{ \frac{B_0(e_h - e_{wh}, v)}{\|v\|_{W^{1,s'}}} \right\} \\ &= \sup_{v \in W^{1,s'}} \left\{ \frac{B_0(e_h, v) - B_1(e_{wh}, v) + \Delta B(e_{wh}, v)}{\|v\|_{W^{1,s'}}} \right\} \\ &= \sup_{v \in W^{1,s'}} \left\{ \frac{\langle \Delta K_0 e_h, v \rangle + \Delta B(e_{wh}, v)}{\|v\|_{W^{1,s'}}} \right\} \\ &\leq C(\|\Delta K_0\|_{\mathcal{L}(W^{-1,s}, W^{-1,s})} \|e_h\|_{W^{1,s}} + \|\Delta B\|_{L^\infty} \|e_{wh}\|_{W^{1,s}}); \end{aligned}$$

from

$$\gamma \|e_{wh}\|_{W^{1,s}} = \gamma \|w_0^h\|_{W^{1,s}} \leq \sup_{v \in W^{1,s'}} \left\{ \frac{B_1(w_0^h, v)}{\|v\|_{W^{1,s'}}} \right\} \leq \|F(u_h, \lambda_0)\|_{W^{-1,s}} \leq C \|e_h\|_{W^{1,s}}$$

and, from Corollary 1.6,

$$\begin{aligned} \|\Delta B\|_{L^\infty} &\leq C \|e_h\|_{W^{1,\infty}} \leq Ch^{\xi-n/\min\{r,p\}} \\ \|\Delta K_0\|_{\mathcal{L}(W^{-1,s}, W^{-1,s})} &\leq C \|e_h\|_{W^{1,\infty}} \leq Ch^{\xi-n/\min\{r,p\}}. \end{aligned}$$

(ii) Similarly, the second inequality is obtained by

$$\begin{aligned} \gamma \|e_{0h} - e_{wh}\|_{W^{1,s}} &\leq \sup_{v \in W^{1,s'}} \left\{ \frac{B_0(e_{0h} - e_{wh}, v)}{\|v\|_{W^{1,s'}}} \right\} \\ &= \sup_{v \in W^{1,s'}} \left\{ \frac{B_0(e_{0h}, v) - B_1(e_{wh}, v) + \Delta B(e_{wh}, v)}{\|v\|_{W^{1,s'}}} \right\} \\ &\leq \sup_{v \in W^{1,s'}} \left\{ \frac{\langle \Delta K_0 e_h, v \rangle + \Delta B(e_{wh}, v)}{\|v\|_{W^{1,s'}}} \right\} \\ &\leq C(\|\Delta R_h\|_{W^{-1,s}} + \|\Delta B\|_{L^\infty} \|e_{wh}\|_{W^{1,s}}) \\ &\leq C(\|\Delta K_0\|_{\mathcal{L}(W^{-1,s}, W^{-1,s})} \|e_h\|_{W^{1,s}} + \|\Delta B\|_{L^\infty} \|e_{wh}\|_{W^{1,s}}). \end{aligned}$$

The rest follows from the previous item.

(iii) From Lemma 2.6 we have that

$$\begin{aligned} \|e_h - e_{0h}\|_{W^{1,s}} &= \|u_h - u_{0h}\|_{W^{1,s}} \leq C \|\Delta K_0\|_{\mathcal{L}(W^{-1,s}, W^{-1,s})} \|e_h\|_{W^{1,s}} \\ &\leq Ch^{\xi-n/\min\{r,p\}} \|e_h\|_{W^{1,s}}. \end{aligned}$$

So, for $h > 0$ small enough, we obtain

$$\|e_h\|_{W^{1,s}} \leq \frac{\|e_{0h}\|_{W^{1,s}}}{1 - Ch^{\xi - n/\min\{r,p\}}}.$$

From Theorem 2.7 and the inequality above we obtain the third and last inequality. Finally we set $\varepsilon = \xi - (n/\min\{r, p\})$. \square

Lemma 2.9. *Let the hypothesis of Theorem 2.7 be satisfied. Furthermore, let $u_0 \in W^{1+\xi,p}$, such that $\xi > n/\min\{r, p\}$. Then, if $\eta_0(s)$, $s \in [r', r]$, is asymptotically exact, so will be $\eta_1(s)$; i.e., if there exists a constant $C \neq C(h)$ and $\varepsilon_0 > 0$, such that, for all small enough $h > 0$,*

$$|\eta_0(s) - \|e_{0h}\|_{W^{1,s}}| \leq Ch^{\varepsilon_0} \|e_{0h}\|_{W^{1,s}},$$

then, the same is true for $\eta_1(s)$, that is, there exists a constant $C \neq C(h)$, and $\varepsilon > 0$, such that

$$|\eta_1(s) - \|e_{wh}\|_{W^{1,s}}| \leq Ch^{\varepsilon} \|e_{wh}\|_{W^{1,s}}.$$

Proof. From Lemma 2.8 there exists $\varepsilon_1 > 0$, such that

$$\begin{aligned} |\eta_1(s) - \|e_{wh}\|_{W^{1,s}}| &\leq |||e_{wh}\|_{W^{1,s}} - \|e_{0h}\|_{W^{1,s}}| + |||e_{0h}\|_{W^{1,s}} - \eta_0(s)| + |\eta_0(s) - \eta_1(s)| \\ &\leq Ch^{\varepsilon_1} \|e_{wh}\|_{W^{1,s}} + Ch^{\varepsilon_1} \|e_{0h}\|_{W^{1,s}} + Ch^{\varepsilon_1} (\|e_{0h}\|_{W^{1,s}} + \eta_0(s)). \end{aligned}$$

Since $\eta_0(s)$ is asymptotically exact

$$\eta_0 \leq (1 + Ch^{\varepsilon_0}) \|e_{0h}\|_{W^{1,s}}.$$

From Lemma 2.8 we get

$$\|e_{0h}\|_{W^{1,s}} \leq (1 + Ch^{\varepsilon_1}) \|e_{wh}\|_{W^{1,s}}.$$

The three inequalities above yield the desired result, by taking $\varepsilon = \min\{\varepsilon_0, \varepsilon_1\}$. \square

Lemma 2.10. *Let the hypotheses of Lemma 2.9 be satisfied for given implicit estimators $\eta_0(s)$ and $\eta_1(s)$, $s \in [r', r]$. Then, for all $h > 0$ small enough, the estimator $\eta_1(s)$ is asymptotically equal to $\|e_h\|_{W^{1,s}}$, for all $s \in [r', r]$. That is, there exists $\varepsilon > 0$ and a constant $C \neq C(h)$, such that*

$$|\eta_1(s) - \|e_h\|_{W^{1,s}}| \leq Ch^{\varepsilon} \|e_h\|_{W^{1,s}},$$

for all $s \in [r', r]$.

Proof. From Lemma 2.8

$$|\eta_1(s) - \|e_h\|_{W^{1,s}}| \leq |\eta_1(s) - \|e_{wh}\|_{W^{1,s}}| + |||e_{wh}\|_{W^{1,s}} - \|e_h\|_{W^{1,s}}|.$$

From Lemma 2.8, we get that there exists $\varepsilon_0 > 0$, such that

$$|||e_{wh}\|_{W^{1,s}} - \|e_h\|_{W^{1,s}}| \leq Ch^{\varepsilon_0} \|e_h\|_{W^{1,s}},$$

and from Lemmas 2.9 and 2.8

$$|\eta_1(s) - \|e_{wh}\|_{W^{1,s}}| \leq Ch^{\varepsilon_1} \|e_{wh}\|_{W^{1,s}} \leq Ch^{\varepsilon_1} (1 + Ch^{\varepsilon_0}) \|e_h\|_{W^{1,s}}.$$

The above inequalities provide the statement of the lemma, by taking $\varepsilon = \min\{\varepsilon_0, \varepsilon_1\}$. \square

Remark. The definition of asymptotic exactness given in the statement of Lemma 2.10 may be weakened by supposing that there exists a function $\mathcal{Q}(h)$, with $\mathcal{Q}(h) \rightarrow 0$ as $h \rightarrow 0$, and such that

$$|\eta_1(s) - \|e_h\|_{W^{1,s}}| \leq \mathcal{Q}(h) \|e_h\|_{W^{1,s}},$$

for all $s \in [r', r]$.

As an example, in order to illustrate the procedures described in this section, we analyse the following partial differential equation, $F : W_0^{1,p_1} \times \mathbb{R} \rightarrow W^{-1,p'_1}$, $1/p_1 + (1/p'_1) = 1$:

$$F(u, \lambda) = -\nabla \cdot [(1 + |\nabla u|^2)^{\frac{p_1-2}{2}} \nabla u] + \lambda u - f.$$

Here we will assume that $p_1 > 1$ and that the domain $\Omega \subset \mathbb{R}^n$ is as smooth as we wish. The above definition implies that $Q(u, \lambda) = -\nabla \cdot [(1 + |\nabla u|^2)^{\frac{p_1-2}{2}} \nabla u]$, $R(u, \lambda) = -\lambda u$ and $f(\lambda) = f$. Furthermore, $A_d = \partial\Omega$ (no Neumann boundary condition). It is observed that when $\lambda > 0$, then $Q - R$ is a uniformly coercive monotone operator and, then, Pr.1 has a unique solution for each such λ . Also, provided u_0 is smooth enough it is not a difficult task to show that Hypothesis 1.2 is satisfied. The smoothness of u_0 depends on the smoothness of $\partial\Omega$ and of f , which are assumed to be as smooth as needed.

Now, we observe that, for all $\psi \in H_0^1$,

$$K_0\psi = D_u F(u_0, \lambda_0)\psi = -\nabla \cdot [A(x) \cdot \nabla \psi] + \lambda_0\psi,$$

where A is the matrix

$$A(x) = (1 + |\nabla u_0|^2)^{(p_1-4)/2} [(p_1 - 2)\nabla u_0 \nabla u_0 + (1 + |\nabla u_0|^2)I].$$

It is easily seen that, for all $\xi \in \mathbb{R}^n$,

$$(A \cdot \xi) \cdot \xi \geq \begin{cases} |\xi|^2 [1 + (p_1 - 1)|\nabla u_0|^2], & 1 < p_1 \leq 2, \\ |\xi|^2 (1 + |\nabla u_0|^2) & p_1 > 2. \end{cases}$$

Then, for all $\lambda_0 > 0$, $D_u F(u_0, \lambda_0) : H_0^1 \rightarrow H^{-1}$ is a uniformly coercive elliptic linear operator with smooth coefficients. If (u_h, λ_0) is the finite element solution to $DPr.1$, which exists following Theorem 1.5, then it converges, following Corollary 1.6, with rate $\min\{q, \xi - n/r\}$, in the $W^{1,\infty}$ -norm, where $q \geq 1$ is the polynomial order of approximation of the shape functions in each element.

In a similar fashion as we did for K_0 , we obtain that, for all $\psi \in H_0^1$

$$K_h\psi = D_u F(u_h, \lambda_0)\psi = -\nabla \cdot [A_h(x) \cdot \nabla \psi] + \lambda_0\psi,$$

where A_h is the matrix

$$A_h(x) = (1 + |\nabla u_h|^2)^{\frac{p_1-4}{2}} [(p_1 - 2)\nabla u_h \nabla u_h + (1 + |\nabla u_h|^2)I].$$

It is easily seen that, for all $\xi \in \mathbb{R}^n$,

$$(A_h \cdot \xi) \cdot \xi \geq \begin{cases} |\xi|^2 [1 + (p_1 - 1)|\nabla u_h|^2], & 1 < p_1 \leq 2, \\ |\xi|^2 (1 + |\nabla u_h|^2) & p_1 > 2, \end{cases}$$

and thus, $K_h : H_0^1 \rightarrow H^{-1}$ is a uniformly coercive linear elliptic operator.

Now, the strategy to obtain $\eta_1(s)$, $s \in [r', r]$ is to estimate e_{wh} , considering the following error equation:

$$B_1(e_{wh}, v) = \langle -F(u_h, \lambda_0), v \rangle \quad \text{for all } v \in W^{1,s'}.$$

Recall that $B_1(\cdot, \cdot) = \langle K_h(\cdot), (\cdot) \rangle$. There are several options for computing the implicit estimator η_1 . For a review of some of them see [7, 16]. The best choices will be among those which may be asymptotically exact, provided some smoothness requirements are satisfied. Particularly, those requirements are met by our abstract and smooth linear problem, defined by the bilinear form B_0 and the right-hand side f_0 . Hence, by Lemma 2.10, $\eta_1(s)$, computed by such a method, will be asymptotically exact with respect to the error $e_h = u_0 - u_h$ in the $W_0^{1,s}$ -norm.

We would like to make two basic concluding remarks. First, the use of auxiliary linear problems for estimating the error for nonlinear problems has been in use for some time. What we have proved is that it is possible to justify the use of such estimators designed for linear problems in nonlinear problems (asymptotically exact implicit estimators preserve that property for the nonlinear problems, provided some standard assumptions are satisfied). Second, the majority of the estimators for linear problems are considered in the norm of some suitable Hilbert space, but a large number of nonlinear problems are not posed on such spaces. Nevertheless, it is our conjecture that the estimators should behave well in other norms, provided some assumptions (e.g., Hypotheses 1.3 and 1.4) about the operator and solution set hold, together with some restrictions on the mesh and on the finite element spaces.

References

- [1] R.A. Adams, Sobolev Spaces, Academic Press, New York, 1975.
- [2] I. Babuška, W.C. Rheinboldt, Computational error estimates and adaptive processes for some nonlinear structural problems, *Comput. Meth. Appl. Mech. Eng.* 34 (1982) 895–937.
- [3] F.C.G. Santos, The linearization method for the numerical analysis of finite element solutions to quasilinear elliptic partial differential equations, Research Report 001-1998, Computational Mechanics Group, UFPE, 1998.
- [4] F.C.G. Santos, A posteriori estimation for quasilinear elliptic partial differential equations, Research Report 002-1998, Computational Mechanics Group, UFPE, 1998.
- [5] I. Babuška, Stroubolis, C.S. Upadhyay, A model study of the quality of A-posteriori estimators for linear elliptic problems: error estimation in the interior of patchwise uniform grids of triangles, Technical Note BN-1147, Institute for Physical Science and Technology, University of Maryland, 1993.
- [6] N.A. Bobylev, Yu.M. Burman, Morse lemma for functionals of variational calculus, *Funktsional'nyi Analiz i Ego Prilozheniya*, 25(3) (1991) 1–11, (English Translation).
- [7] G. Caloz, J. Rappaz, Numerical analysis for nonlinear and bifurcation problems, To appear in: L.J. Lions, (Eds.), *Handbook of Numerical Analysis*, P.G. Ciarlet, North-Holland, Amsterdam.
- [8] M. Crouzeix, J. Rappaz, On Numerical Approximation in Bifurcation Theory, RMA 13, Masson, Paris, 1989.
- [9] G.M. Krasnosel'skii, G.M. Vainikko, P.P. Zabreiko, Ya. B. Rutitskii, V.Ya. Stetsenko, Approximate Solutions of Operator Equations, Wolters-Noordhoff, Groninger, 1972.

- [10] J. Poussin, J. Rappaz, Consistency, stability, A-priori and A-posteriori errors for Petrov-Galerkin methods applied to nonlinear problems, Nr. 04.92, École Polytechnique Fédérale de Lausanne, 1992.
- [11] W.C. Rheinboldt, Numerical Analysis of Parameterized Nonlinear Equations, Wiley, New York, 1986.
- [12] F.C.G. Santos, Numerical analysis of parameterized partial differential equations, Ph.D. Thesis, University of Maryland, 1995.
- [13] T. Tsuchiya, A-priori and A-posteriori error estimates of finite element solutions of parameterized nonlinear equations, Ph.D. Thesis, University of Maryland, 1990.
- [14] R. Verfürth, A review of A-posteriori error estimation and adaptive mesh refinement, Tech. Rep., Institute Für Angewandte Mathematik, Universität Zürich, 1993.